Report on **Bachelor Thesis**

Institute of Economic Studies, Faculty of Social Sciences, Charles University

Student:	**Vojtěch Chalupa**
Advisor:	**Matěj Bajgar, DPhil.**
Title of the thesis:	**Determinants of football players' market value during the season: Discrepancies between TOP 5 and the rest of the leagues**

OVERALL ASSESSMENT (provided in English, Czech, or Slovak):

Please provide a short summary of the thesis, your assessment of each of the four key categories, and an overall evaluation and suggested questions for the discussion. The minimum length of the report is 300 words.

Short summary
The thesis studies determinants the determinants of football players' market values. It applies OLS and LASSO regressions to large player-level data for the 2022/2023 season, compiled from several different sources, to estimate how players' value is related to player and team characteristics. It finds that there are substantial differences between which characteristics predict players' value in the top 5 leagues and in other competition. In particular, statistics associated with players holding a ball a lot are positive predictors of value in the lower-ranked competitions but not in the top ones. The thesis also finds a consistently negative effects of players' age, which is in contrast to previous literature.

Contribution
The thesis clearly explains its contribution to existing literature on player valuation.
1) While previous studies often focused on a particular league, the thesis compares many leagues and contrasts the top 5 leagues with the lower-ranked ones.
2) Where previous studies often focused on a particular position, the thesis compares different player positions.

The contribution of the thesis is not outstanding, but it is well explained and completely adequate for a bachelor's thesis.

Methods
To measure players' market value, the paper uses crowd-sourced estimates from the Transfermarkt.com webiste. It compelements the market value data with players characteristics from the Football Manager 23 video game database and match data from Fbref.com.

The paper estimates the effect of player and team characteristics on player value using standard OLS regressions. It also uses the LASSO methodology to see if the list of explanatory variables could be narrowed down, and it drops two explanatory variables as a result.

The overall methodology is quite basic but adequate for this type of work. It is correctly applied and described.

That said, I see some potential issues with the approach used that should be discussed during the defense.

1) <u>The measure of market value.</u> The outcome measure in the analysis is players' market value estimated by the fans on Transfermarkt.com. I understand it is convenient that these estimates are available for all players (while only a minority of players are traded in any given year), and the paper discusses evidence suggesting that these estimates are quite close to the actual transfer fees when transfers take place. However, it should be acknowledged more clearly that the actual market value is observed only during the transfers and the used estimates are just proxies. In particular, I expected the paper would conduct a robustness analysis that would narrow down the

Report on Bachelor Thesis

Institute of Economic Studies, Faculty of Social Sciences, Charles University

Student:	**Vojtěch Chalupa**
Advisor:	**Matěj Bajgar, DPhil.**
Title of the thesis:	**Determinants of football players' market value during the season: Discrepancies between TOP 5 and the rest of the leagues**

sample to players who had a transfer during the season and whose actual market value is thus known. Instead, to my surprise, I learnt that not only is not such robustness analysis not carried out, but such players are completely dropped from the analysis (!): *„Furthermore, we exclude players who were transferred to a new club either during the summer or winter trading window. The primary concern is that market values are typically updated at the beginning and after the end of transfer periods. In the case of transfers, the player's value during the updates may be significantly influenced by the transfer fee which can result in a substantial increase in the player's market value without sufficient performance-based justification."* As an economist, I find it hardly unusual to treat actual market transactions where clubs put real money on the table (a lot of it!) as just some kind of noise that blurs our vision of the Platonian „true" player value, as estimated by the fans. If there are practical reasons for relying on third-party estimates over actual market prices (e.g. transfer values are not available in available data) this should be clearly stated and the approach then justified as „second-best". Or if there is a particular interest in studying fan opinions over clubs', the results should be interpreted in that light.

2) <u>One-year sample.</u> Is there a particular reason for limiting the sample to a single year? Is that because only contemporary info is available in the data? This should be clarified because, in general, using panel data would allow for a much richer analysis that could, among other things, control for time-invariant unobservable player characteristics.

3) <u>Interpretation of the lagged dependent variable.</u> The model includes the pre-season market value as one of the explanatory variables. The discussion of the results highlights that this variable comes out the strongest in the results, but the variable is otherwise discussent like any other explanatory variable. What I am missing is a more thoughtful interpretation of this variable and a discussion of how its inclusion changes the interpretation of the model, turning it into a dynamic specification where market value adjusts based on the characteristics (e.g. goals scored) only gradually over time, so short-term and long-term effects can be separately expressed. Is such specification justified in the given context? Should we interpret the model as studying *updates* in the value based on recent performance? But shouldn't players' recent performance then be measured relative to their previous performance? For some players, scoring 5 goals in a season would be a positive surprise and greatly increase their value. For Kylian Mbappe, it would be a massive disappointment and his value would crater.

4) <u>Elasticity interpretation of log specification.</u> The paper says that *„a one-unit change in an independent variable results in a —beta*100% modification in MVA."* This is a minor point, but it should have been noted that the proportional effect should be calculated as exp(beta)-1, and this will be similar to beta only for small changes in the explanatory variables and small estimates, but for bigger changes it can be quite different.

Literature
The thesis shows a good command of the related literature, accurately sets the thesis within the context this literature and clearly describes the contribution of the thesis. Studies are discussed with clarity and appropriately cited.

Manuscript form
The thesis is generally clear and well written. The thesis is well structured and appropriately formatted in LaTex.

Report on **Bachelor Thesis**

Institute of Economic Studies, Faculty of Social Sciences, Charles University

Student:	**Vojtěch Chalupa**
Advisor:	**Matěj Bajgar, DPhil.**
Title of the thesis:	**Determinants of football players' market value during the season: Discrepancies between TOP 5 and the rest of the leagues**

Overall evaluation and suggested questions for the discussion during the defense

This is generally a well executed thesis that certainly meets the criteria for a bachelor's thesis at the IES. That said, in terms of originality and methodology it is not at the same level as the very best theses at the IES. For these reasons I suggest grade B or C based on the performance during the defense.

During the defense, the committee may want to ask the following questions:

1. What is your rationale for using fan estimates rather than actual transfer fees, and in particular why are you dropping players who had a transfer during the season? Should these players have the estimates closest to the actual market prices?
2. The sample contains only about 12 players per team. Even excluding goalkeepers, the teams in these top competition certainly have many more players who get to play over 90 minutes over the course of a seasons, and I also assume players on loans are not very common. What is then the reason for these low numbers? Is it the exclusion of players who had a transfer? Or are there gaps in the data?
3. How should you estimates be interpreted given that you included a lagged dependent variable? What leads you to their inclusion? Could you estimate a fixed effects or first differenced model instead?
4. What is the reason for looking only at a single year of data?

SUMMARY OF POINTS AWARDED *(for details, see below):*

CATEGORY		POINTS
Contribution	*(max. 30 points)*	20
Methods	*(max. 30 points)*	20
Literature	*(max. 20 points)*	20
Manuscript Form	*(max. 20 points)*	20
TOTAL POINTS	*(max. 100 points)*	**80**
GRADE	**(A – B – C – D – E – F)**	**C**

NAME OF THE REFEREE: *Matěj Bajgar, DPhil.*

DATE OF EVALUATION: *4th June 2024*

_____

Referee Signature

Report on **Bachelor Thesis**

Institute of Economic Studies, Faculty of Social Sciences, Charles University

Student:	**Vojtěch Chalupa**
Advisor:	**Matěj Bajgar, DPhil.**
Title of the thesis:	**Determinants of football players' market value during the season: Discrepancies between TOP 5 and the rest of the leagues**

**EXPLANATION OF CATEGORIES AND SCALE:**

CONTRIBUTION: _The author presents original ideas on the topic demonstrating critical thinking and ability to draw conclusions based on the knowledge of relevant theory and empirics. There is a distinct value added of the thesis._

METHODS: _The tools used are relevant to the research question being investigated, and adequate to the author's level of studies. The thesis topic is comprehensively analyzed._

LITERATURE REVIEW: _The thesis demonstrates author's full understanding and command of recent literature. The author quotes relevant literature in a proper way._

MANUSCRIPT FORM: _The thesis is well structured. The student uses appropriate language and style, including academic format for graphs and tables. The text effectively refers to graphs and tables and disposes with a complete bibliography._

**Overall grading:**

TOTAL	GRADE
91 – 100	**A**
81 - 90	**B**
71 - 80	**C**
61 – 70	**D**
51 – 60	**E**
0 – 50	**F**